Journal of the Operations Research
Society of Japan
Vol. 39, No. 1, March 1996

A PRACTICAL ALGORITHM FOR MINIMIZING A RANK-TWO SADDLE FUNCTION ON A POLYTOPE

Takahito Kuno
University of Tsukuba

(Received February 25, 1994; Final December 5, 1994)

Abstract This paper addresses a practical method for minimizing a class of saddle functions $f : \mathbf{R}^n \rightarrow \mathbf{R}^1$ on a polytope. Function f is continuous and possesses a rank-two property, i.e., the value of f is defined only by two linearly independent vectors. It is shown that a parametric right-hand-side simplex algorithm decomposes the problem into a finite sequence of one-dimensional subproblems. A globally ϵ-optimal solution of each subproblem is obtained by using a successive underestimation method. Computational results indicate that the algorithm can solve fairly large scale problems efficiently.

1. Introduction

In this paper we will develop a practical algorithm for minimizing a class of saddle functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$, i.e.,

$$\text{minimize}\{f(x) \mid x \in D\}, \tag{1.1}$$

where $D \subset \mathbb{R}^n$ is a polytope. We assume that f is continuous and possesses the *rank-two property* with respect to two linearly independent vectors $c_1, c_2 \in \mathbb{R}^n$. This means that there exists a continuous function $g : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ such that $f(x) = g({c_1}^T x, {c_2}^T x)$ for all $x \in \mathbb{R}^n$ [15], though we need not know g explicitly in our algorithm. Since f is a saddle function, $g(\cdot, {c_2}^T x)$ and $g({c_1}^T x, \cdot)$ are convex and (quasi)concave functions respectively for any fixed $x \in \mathbb{R}^n$. Due to this *convex-concave property* of f, there are multiple locally optimal solutions in D. In contrast to (quasi)concave minimization problems, (1.1) might have no globally optimal solutions among vertices of D.

Saddle functions are well known in many literature in the context of minimax problems. In [17] Muu and Oettli have solved a more general class of (1.1), in which f is a full-rank saddle function. Muu has also considered a problem containing a full-rank saddle function in the constraint rather than in the objective function [16]. However, the algorithms developed for the general purpose can usually handle only instances of a very limited scale. We will therefore exploit the rank-two property of f and show that a parametric simplex algorithm decomposes (1.1) into a finite sequence of one-dimensional subproblems, which can be solved very efficiently.

Rank-two nonconvex minimization problems are important in practical applications such as bicriterion decision making [4, 8], computational geometry [11, 14] or network flow problems [25] to name only a few (see [24]). Many of them, involving linear multiplicative programs [9, 18, 23] and certain d.c. programs (minimizations of the difference of two convex functions $h_1({c_1}^T x) - h_2({c_2}^T x)$) [22], belong to the class (1.1).

The author was partially supported by Grand-in-Aid for Scientific Research of the Ministry of Education, Science and Culture, Grant No. (C)05650061.

In Section 2 we will show that (1.1) can be solved by solving a sequence of one-dimensional problems of the same form as (1.1). The sequence can be generated by applying a parametric right-hand-side simplex algorithm to two linear programs associated with (1.1). Section 3 is devoted to the procedure for obtaining a globally ϵ-optimal solution of one-dimensional problems. By exploiting the convex-concave property of f we will construct a branch-and-bound algorithm based on a successive underestimation method [7]. Results of computational experiment on the algorithm are presented in Section 4. In Section 5 we will briefly discuss the average performance of the algorithm when we apply it to certain nonconvex quadratic programs.

2. Decomposition of the Problem into One-Dimensional Problems

The problem we consider in this paper is as follows:

$$(\mathrm{P}) \quad \left| \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b, \ x \geq 0, \end{array} \right.$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $f : \mathbb{R}^n \to \mathbb{R}^1$ is a continuous function. There are two linearly independent vectors $c_1, c_2 \in \mathbb{R}^n$ which characterize f. Namely,

(i) Rank-two property: For any $x \in \mathbb{R}^n$

$$d \in \mathbb{R}^n, \ {c_k}^T d = 0, \ k = 1, 2 \Longrightarrow f(x + d) = f(x). \tag{2.1}$$

(ii) Convex-concave property: For any $x \in \mathbb{R}^n$

$$d \in \mathbb{R}^n, \ {c_2}^T d = 0 \Longrightarrow f(x + \lambda d) \leq (1 - \lambda) f(x) + \lambda f(x + d), \ \forall \lambda \in [0, 1], \tag{2.2}$$

$$d \in \mathbb{R}^n, \ {c_1}^T d = 0 \Longrightarrow f(x + \lambda d) \geq \min\{f(x), f(x + d)\}, \ \forall \lambda \in [0, 1]. \tag{2.3}$$

We assume in the sequel that the feasible region:

$$D = \{x \in \mathbb{R}^n \mid Ax = b, \ x \geq 0\}. \tag{2.4}$$

is nonempty and bounded, which implies that (P) has a globally optimal solution. Figure 2.1 shows a two-dimensional example, where $c_1 = (0, 1)^T$, $c_2 = (1, 0)^T$ and $f(x) = -x_1^2 + x_2^2$. It is easy to see that this function has three local minimum points A, B and C, among which C is the global one.

Let $\zeta = {c_1}^T x$ for an arbitrary $x \in D$ and consider a subproblem of (P):

$$(\mathrm{P}(\zeta)) \quad \left| \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in D, \ {c_1}^T x = \zeta. \end{array} \right.$$

Then $(\mathrm{P}(\zeta))$ is feasible and has an optimal solution which coincides with that of a linear program, i.e., either

$$(\mathrm{PL}_1(\zeta)) \left| \begin{array}{ll} \text{minimize} & {c_2}^T x \\ \text{subject to} & x \in D, \ {c_1}^T x = \zeta, \end{array} \right.$$

or

$$(\mathrm{PL}_2(\zeta)) \left| \begin{array}{ll} \text{maximize} & {c_2}^T x \\ \text{subject to} & x \in D, \ {c_1}^T x = \zeta. \end{array} \right.$$

Let $x^k(\zeta)$ be an optimal solution of $(\mathrm{PL}_k(\zeta))$ $(k = 1, 2)$ and define

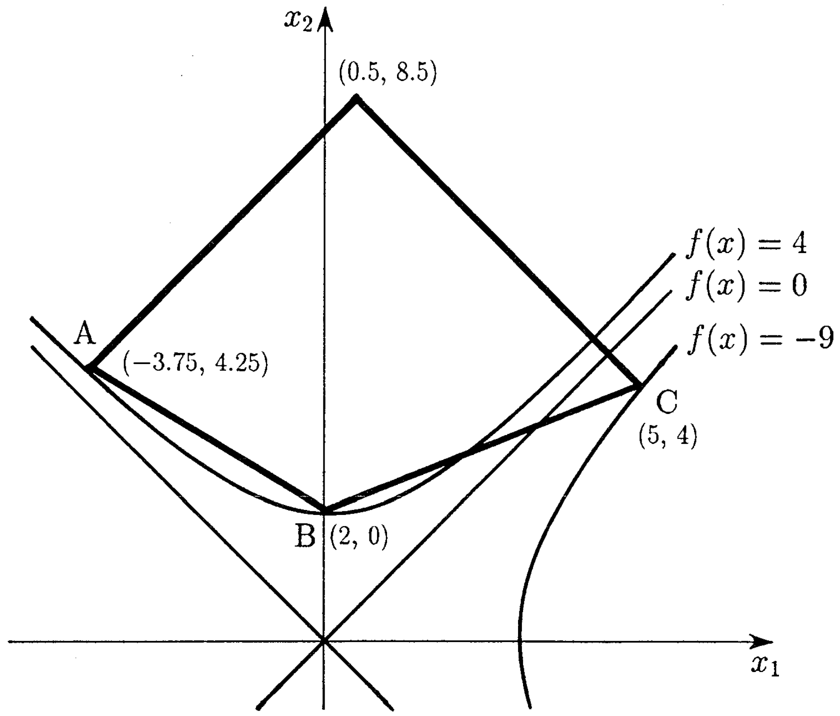

Figure 2.1. Example of (P) in R^2.

$$x^*(\zeta) \in \operatorname{argmin}\{f(x) \mid x = x^k(\zeta),\ k = 1,\ 2\}.$$

Lemma 2.1. *If $\zeta = {c_1}^T x$ for some $x \in D$, then $x^*(\zeta)$ is optimal to $(P(\zeta))$.*

Proof: By the rank-two property, f is a function of a single variable $\eta = {c_2}^T x$ if the value ${c_1}^T x$ is fixed at ζ. The values ${c_2}^T x^1(\zeta)$ and ${c_2}^T x^2(\zeta)$ are the minimum and the maximum of η respectively. It follows from (2.3) of property (ii) that the minimum of f is attained at either of the extreme points of the interval $[{c_2}^T x^1(\zeta),\ {c_2}^T x^2(\zeta)]$. □

Let

$$\zeta_{\min} = \min\{{c_1}^T x \mid x \in D\};\quad \zeta_{\max} = \max\{{c_1}^T x \mid x \in D\}.$$

It is obvious that a globally optimal solution of (P) can be obtained by solving $(P(\zeta))$ for all $\zeta \in [\zeta_{\min},\ \zeta_{\max}]$. By Lemma 2.1, this can be done if we solve the two linear programs $(PL_1(\zeta))$ and $(PL_2(\zeta))$ as varying the value ζ over the interval $[\zeta_{\min},\ \zeta_{\max}]$.

Theorem 2.2. *There exists $\zeta \in [\zeta_{\min},\ \zeta_{\max}]$ such that $x^*(\zeta)$ is a globally optimal solution of (P).* □

Let us apply a parametric right-hand-side simplex method (abbreviated as PRSM) to $(PL_k(\zeta))$ $(k = 1, 2)$. For the sake of simplicity we impose here the dual nondegeneracy assumption:

Assumption 2.1. *Both $(PL_1(\zeta))$ and $(PL_2(\zeta))$ have a unique optimal solution for any $\zeta \in [\zeta_{\min},\ \zeta_{\max}]$.* □

Suppose we have an optimal basis $B_0^k \in \mathbb{R}^{(m+1)\times(m+1)}$ of $(\mathrm{PL}_k(\zeta_{\min}))$. Under the above assumption B_0^k remains optimal even if the value of ζ slightly increases from $\zeta_{\min}$. However, when ζ is beyond some point, say ζ_1^k $(\leq \zeta_{\max})$, some basic variable turns negative and the primal feasibility of B_0^k is violated. Then we carry out a single dual pivot replacing the basic variable with an appropriate nonbasic one and obtain an alternative basis B_1^k, which is optimal to $(\mathrm{PL}_k(\zeta_1^k))$ (see e.g. [2, 3] for further details).

Applying these operations iteratively, we can generate a sequence of subintervals $[\zeta_0^k, \zeta_1^k]$, $[\zeta_1^k, \zeta_2^k], \ldots, [\zeta_{p_k-1}^k, \zeta_{p_k}^k]$ in the interval $[\zeta_{\min}, \zeta_{\max}]$, where $\zeta_0^k = \zeta_{\min}$, $\zeta_{p_k}^k = \zeta_{\max}$ and $\zeta_{i+1}^k > \zeta_i^k$ for each i. Simultaneously, we have the associated sequence of bases $B_0^k, B_1^k, \ldots, B_{p_k-1}^k \in \mathbb{R}^{(m+1)\times(m+1)}$ such that B_i^k is optimal to $(\mathrm{PL}_k(\zeta))$ for all $\zeta \in [\zeta_i^k, \zeta_{i+1}^k]$. We denote $[\zeta_i^k, \zeta_{i+1}^k]$ by Z_i^k in the sequel. As well known, $x^k(\zeta)$ is an affine function over each Z_i^k and can be expressed as

$$x^k(\zeta) = \frac{\zeta_{i+1}^k - \zeta}{\zeta_{i+1}^k - \zeta_i^k}\, x^k(\zeta_i^k) + \frac{\zeta - \zeta_i^k}{\zeta_{i+1}^k - \zeta_i^k}\, x^k(\zeta_{i+1}^k), \quad \zeta \in \mathrm{Z}_i^k.$$

If for every i we can compute

$$x^k(\mathrm{Z}_i^k) \in \operatorname{argmin}\{f(x) \mid x = (1-\lambda)x^k(\zeta_i^k) + \lambda x^k(\zeta_{i+1}^k),\ \lambda \in [0, 1]\},$$

then Theorem 2.2 guarantees that

$$x^* \in \operatorname{argmin}\{f(x) \mid x = x^k(\mathrm{Z}_i^k),\ j = 0, 1, \ldots, p_k - 1,\ k = 1, 2\}$$

is a globally optimal solution of (P). The procedure for computing $x^k(\mathrm{Z}_i^k)$ will be presented in the next section.

We summarize the algorithm below:

Algorithm PRSM.

Step 1. Solve a linear program: minimize$\{{c_1}^T x \mid x \in D\}$ and obtain an optimal basis B° and the associated optimal solution x°. Initialize the incumbent: $x^* = x^\circ$, $v^* = f(x^*)$. Let $k = 1$ and go to Step 2.

Step 2. Let $\underline{\zeta} = {c_1}^T x^\circ$ and $\underline{B} = B^\circ$. Solve a linear program $(\mathrm{PL}_k(\zeta))$ parametrically by increasing ζ from $\underline{\zeta}$:

1° If $(\mathrm{PL}_k(\zeta))$ is infeasible for $\zeta > \underline{\zeta}$, then go to Step 3.

2° Determine a value $\overline{\zeta}$ of ζ such that $\underline{B}$ is an optimal basis for all $\zeta \in \mathrm{Z} = [\underline{\zeta}, \overline{\zeta}]$. Using a dual pivot operation, obtain an alternative basis $\overline{B}$ which is optimal to $(\mathrm{PL}_k(\overline{\zeta}))$.

3° Compute $x^k(\mathrm{Z}) \in \operatorname{argmin}\{f(x) \mid x = (1-\lambda)x^k(\underline{\zeta}) + \lambda x^k(\overline{\zeta}),\ \lambda \in [0, 1]\}$. If $f(x^k(\mathrm{Z})) < v^*$, then update the incumbent: $x^* = x^k(\mathrm{Z})$, $v^* = f(x^*)$.

4° Let $\underline{\zeta} = \overline{\zeta}$, $\underline{B} = \overline{B}$ and go to 1°.

Step 3. If $k = 2$, then terminate. Otherwise, let $k = 2$ and go to Step 2. □

Under Assumption 2.1, the above algorithm terminates after finitely many iterations yielding an optimal solution x^* of (P) if Step 2. 3° can be done in finite time. In the case of degeneracy, we have to use a suitable pivoting rule to avoid cycling (see e.g. [2]).

3. Successive Underestimation Method for One-Dimensional Problems

In this section we consider the problem to be solved in Step 2. 3° of algorithm PRSM, i.e., for each $k = 1, 2$,

$$(\mathrm{P}_k(\mathrm{Z}))\ \left|\ \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x = (1-\lambda)x^k(\underline{\zeta}) + \lambda x^k(\overline{\zeta}),\ \lambda \in [0, 1], \end{array}\right.$$

where $x^k(\underline{\zeta})$ and $x^k(\overline{\zeta})$ are optimal solutions of $(\mathrm{PL}_k(\underline{\zeta}))$ and $(\mathrm{PL}_k(\overline{\zeta}))$ respectively, and $\mathrm{Z} = [\underline{\zeta}, \overline{\zeta}]$ is a subinterval of $[\zeta_{\min}, \zeta_{\max}]$ such that a basis $\underline{B}$ is optimal to $(\mathrm{PL}^k(\zeta))$ for all $\zeta \in \mathrm{Z}$. The difference between $(\mathrm{P}_k(\mathrm{Z}))$ and (P) is that the feasible region of the former:

$$D_k(\mathrm{Z}) = \{x \in \mathbb{R}^n \mid x = (1-\lambda)x^k(\underline{\zeta}) + \lambda x^k(\overline{\zeta}),\ \lambda \in [0, 1]\}$$

is only a line segment. Hence, if f is either convex or concave over $D_k(\mathrm{Z})$, we can compute a minimum $x^k(\mathrm{Z})$ very efficiently by using any one of ordinary methods. This involves the case in which either ${c_2}^T x$ or ${c_1}^T x$ is a constant for any $x \in D_k(\mathrm{Z})$. Although both the values are affine functions of λ over $D_k(\mathrm{Z})$, they are not constants in general. We will therefore propose a successive underestimation method for obtaining a globally ϵ-optimal solution of $(\mathrm{P}_k(\mathrm{Z}))$.

3.1. Lower bounds of the objective function value

We first define a vector $\tilde{c}_1 \in \mathbb{R}^n$ below:

$$\tilde{c}_1 = c_1 - ({c_1}^T c_2 \,/\, \|c_2\|^2)c_2. \tag{3.1}$$

Then we have ${c_1}^T \tilde{c}_1 > 0$ and ${c_2}^T \tilde{c}_1 = 0$ by noting that c_1 and c_2 are linearly independent. Hence by (2.2) of property (ii) function f is convex with respect to the direction $\tilde{c}_1$. We can compute the following by using convex minimization:

$$\underline{L}_k(\mathrm{Z}) = \operatorname{argmin}\{f(x) \mid x = x^k(\underline{\zeta}) + \lambda\alpha(\mathrm{Z})\tilde{c}_1,\ \lambda \in [0, 1]\}, \tag{3.2}$$

$$\overline{L}_k(\mathrm{Z}) = \operatorname{argmin}\{f(x) \mid x = x^k(\overline{\zeta}) - \lambda\alpha(\mathrm{Z})\tilde{c}_1,\ \lambda \in [0, 1]\}, \tag{3.3}$$

where

$$\alpha(\mathrm{Z}) = (\overline{\zeta} - \underline{\zeta}) \,/\, ({c_1}^T \tilde{c}_1). \tag{3.4}$$

Let

$$v^k(\mathrm{Z}) = \min\{f(x) \mid x \in \underline{L}_k(\mathrm{Z}) \cup \overline{L}_k(\mathrm{Z})\}. \tag{3.5}$$

Lemma 3.1. *For any subinterval* $\mathrm{Z}' = [\underline{\zeta}', \overline{\zeta}'] \subset \mathrm{Z}$ *the following relationship holds:*

$$v^k(\mathrm{Z}) \le v^k(\mathrm{Z}').$$

Proof: Choose an arbitrary $x' \in \underline{L}_k(\mathrm{Z}')$. Then there exists $\lambda' \in [0, 1]$ such that $x' = x^k(\underline{\zeta}') + \lambda'\alpha(\mathrm{Z}')\tilde{c}_1$. By linearity of $x^k(\zeta)$ over Z and definition of $\alpha(\mathrm{Z})$ we have

$$x^k(\underline{\zeta}') = \frac{\overline{\zeta} - \underline{\zeta}'}{\overline{\zeta} - \underline{\zeta}} x^k(\underline{\zeta}) + \frac{\underline{\zeta}' - \underline{\zeta}}{\overline{\zeta} - \underline{\zeta}} x^k(\overline{\zeta});\quad \alpha(\mathrm{Z}') = \frac{\overline{\zeta}' - \underline{\zeta}'}{\overline{\zeta} - \underline{\zeta}} \alpha(\mathrm{Z}).$$

Hence x' is written as

$$x' = (1-\beta)x^k(\underline{\zeta}) + \beta x^k(\overline{\zeta}) + \lambda'\gamma\alpha(\mathrm{Z})\tilde{c}_1, \tag{3.6}$$

where $\beta = (\zeta' - \underline{\zeta}) / (\overline{\zeta} - \underline{\zeta})$ and $\gamma = (\overline{\zeta}' - \underline{\zeta}') / (\overline{\zeta} - \underline{\zeta})$. Let

$$\underline{x} = x^k(\underline{\zeta}) + (\beta + \lambda'\gamma)\alpha(\mathrm{Z})\tilde{c}_1; \quad \overline{x} = x^k(\overline{\zeta}) - (1 - \beta - \lambda'\gamma)\alpha(\mathrm{Z})\tilde{c}_1.$$

Then (3.6) is reduced to the following:

$$x' = (1 - \beta)\underline{x} + \beta\overline{x}.$$

Since ${c_1}^T x^k(\underline{\zeta}) = \underline{\zeta}$ and ${c_1}^T x^k(\overline{\zeta}) = \overline{\zeta}$ by definition, we see that

$${c_1}^T(\overline{x} - \underline{x}) = {c_1}^T x^k(\overline{\zeta}) - {c_1}^T x^k(\underline{\zeta}) - \alpha(\mathrm{Z}){c_1}^T\tilde{c}_1 = \overline{\zeta} - \underline{\zeta} - (\overline{\zeta} - \underline{\zeta}) = 0$$

by noting (3.4). We can also check that $\beta + \gamma\lambda' \in [0, 1]$ if $\mathrm{Z}' \subset \mathrm{Z}$. Hence by (2.3) of property (ii) and definition of $v^k(\mathrm{Z})$ we obtain

$$f(x') \geq \min\{f(\underline{x}), f(\overline{x})\} \geq v^k(\mathrm{Z}).$$

Similarly, we have $f(x) \geq v^k(Z)$ for any $x \in \overline{L}_k(\mathrm{Z}')$. □

As a corollary of this lemma, we can show that $v^k(\mathrm{Z})$ gives a lower bound of the optimal value $f(x^k(\mathrm{Z}))$ of problem $(\mathrm{P}_k(\mathrm{Z}))$:

Lemma 3.2. *For any $x \in D_k(\mathrm{Z})$ the following holds:*

$$f(x) \geq v^k(\mathrm{Z}). \tag{3.7}$$

Proof: For any $x \in D_k(\mathrm{Z})$ there exists some $\zeta' \in \mathrm{Z}$ such that $x = x^k(\zeta')$. Hence (3.7) is derived by applying Lemma 3.1 to $\mathrm{Z}' = [\zeta', \zeta'] \subset \mathrm{Z}$. □

Note that $v^k(\mathrm{Z})$ is the optimal value of a relaxed problem of $(\mathrm{P}_k(\mathrm{Z}))$:

$$\left|\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in R\ (\supset D_k(\mathrm{Z})), \end{array}\right.$$

where R is a rectangle with vertices $x^k(\underline{\zeta})$ and $x^k(\overline{\zeta})$ in the plane spanned by c_1 and c_2. Each side of R is collinear with either $\tilde{c}_1$ or c_2, and hence by quasiconcavity of f with respect to c_2 the minimum is achieved at some points, say $\underline{L}_k(\mathrm{Z})$ or $\overline{L}_k(\mathrm{Z})$, on the sides collinear with $\tilde{c}_1$.

Lemma 3.2 enables us to discard $D_k(\mathrm{Z})$ in the course of locating a globally optimal solution of (P) in PRSM when

$$v^k(\mathrm{Z}) \geq f(x^*) \tag{3.8}$$

holds for the best feasible solution x^* obtained by that time. In this case we cannot update the incumbent better than x^* by any point of $D_k(\mathrm{Z})$.

3.2. Branch-and-bound procedure

Let us suppose that (3.8) does not hold. When some point x' in the set:

$$L_k(\mathrm{Z}) = \{x \in \underline{L}^k(\mathrm{Z}) \cup \overline{L}^k(\mathrm{Z}) \mid f(x) = v^k(\mathrm{Z})\} \tag{3.9}$$

is found to be a feasible solution of (P), we may discard $D_k(\mathrm{Z})$ and proceed to the next step after revising the incumbent x^* by x'. If such an x' cannot be found, i.e., $L_k(\mathrm{Z}) \cap D = \emptyset$, we have to search $D_k(\mathrm{Z})$ for a better feasible solution than x^*.

Let us bisect the interval $\mathrm{Z} = [\underline{\zeta}, \overline{\zeta}]$ into $\mathrm{Z}_{11} = [\underline{\zeta}, \zeta_0]$ and $\mathrm{Z}_{12} = [\zeta_0, \overline{\zeta}]$, where $\zeta_0 = (\underline{\zeta} + \overline{\zeta}) / 2$. Then the value $f(x^k(\zeta_0))$ is an upper bound of the optimal value of $(\mathrm{P}_k(\mathrm{Z}))$. If $f(x^k(\zeta_0)) < f(x^*)$, then we need to update the incumbent as $x^* = x^k(\zeta_0)$. Note that we can compute $x^k(\zeta_0)$ without performing any pivot operations, since $x^k(\zeta)$ is affine over the interval Z. We next construct the problems $(\mathrm{P}_k(\mathrm{Z}_{11}))$ and $(\mathrm{P}_k(\mathrm{Z}_{12}))$ associated with the intervals Z_{11} and Z_{12} respectively, and compute lower bounds $v^k(\mathrm{Z}_{11})$ and $v^k(\mathrm{Z}_{12})$ of their optimal values. It is obvious that $D_k(\mathrm{Z}_{11}) \cup D_k(\mathrm{Z}_{12}) = D_k(\mathrm{Z})$ and $D_k(\mathrm{Z}_{11}) \cap D_k(\mathrm{Z}_{12}) = \{x^k(\zeta_0)\}$. Let us define a piecewise constant function on $D_k(\mathrm{Z})$:

$$g_1(x) = \begin{cases} v^k(\mathrm{Z}_{11}), & x \in D_k(\mathrm{Z}_{11}), \\ v^k(\mathrm{Z}_{12}), & x \in D_k(\mathrm{Z}) \setminus D_k(\mathrm{Z}_{11}). \end{cases}$$

Then by Lemma 3.1 we see that

$$v^k(\mathrm{Z}) \leq g_1(x) \leq f(x), \quad \forall x \in D_k(\mathrm{Z}).$$

A further bisection of $\mathrm{Z}_{1\ell}$ with $v^k(\mathrm{Z}_{1\ell}) = \min\{v^k(\mathrm{Z}_{11}), v^k(\mathrm{Z}_{12})\}$ at its middle point ζ_1 can generate an alternative function g_2, which underestimates f over $D_k(\mathrm{Z})$ more exactly than g_1.

If we iterate the above operations as selecting one subinterval of Z giving the least lower bound among them, we will obtain a sequence of piecewise constant functions g_j's such that

$$(v^k(\mathrm{Z}) =)\ g_0(x) \leq g_1(x) \leq g_2(x) \leq \cdots \leq f(x), \quad \forall x \in D_k(\mathrm{Z}).$$

Note that $x^k(\zeta_j)$ is a minimizer of g_j and a jumping point of g_{j+1}. The incumbent x^* is updated by $x^k(\zeta_j)$ when necessary. If

$$g_q(x^k(\zeta_q)) \geq f(x^*)$$

happens to hold, then two cases are possible: (i) x^* is an optimal solution of $(\mathrm{P}_k(\mathrm{Z}))$ if $x^* \in D_k(\mathrm{Z})$, (ii) there are no globally optimal solution of (P) in $D_k(\mathrm{Z})$ otherwise. In either case we can terminate the procedure. Figure 3.1 illustrates the procedure when we apply it to the example shown in Section 2. Here $D_k(\mathrm{Z})$ corresponds to the edge B-C of D in Figure 2.1.

The procedure is summarized as the following branch-and-bound algorithm. Here $\epsilon \geq 0$ is a given tolerance, x^* and v^* are the incumbent and its objective function value respectively.

Procedure BBP(k, x^*, v^*, Z).

1° Compute $v^k(\mathrm{Z})$ and $L_k(\mathrm{Z})$ according to (3.1) – (3.5) and (3.9). If $v^k(\mathrm{Z}) \geq f(x^*)$, then terminate. Otherwise, let $\mathcal{Z} = \{\mathrm{Z}\}$ and $j = 0$.

2° Select an interval $\mathrm{Z}_j = [\underline{\zeta}_j, \overline{\zeta}_j] \in \mathcal{Z}$ with the least $v^k(\mathrm{Z}_j)$ and let $\mathcal{Z} = \mathcal{Z} \setminus \{\mathrm{Z}_j\}$. If $L_k(\mathrm{Z}_j) \cap D \neq \emptyset$, then terminate after revising the incumbent: $x^* = x'$, $v^* = f(x^*)$ for an arbitrary $x' \in L_k(\mathrm{Z}_j) \cap D$.

3° Let $\zeta_j = (\underline{\zeta}_j + \overline{\zeta}_j) / 2$. If $f(x^k(\zeta_j)) < v^*$, then update the incumbent: $x^* = x^k(\zeta_j)$, $v^* = f(x^*)$. If

$$f(x^*) - v^k(\mathrm{Z}_j) \leq \epsilon, \tag{3.10}$$

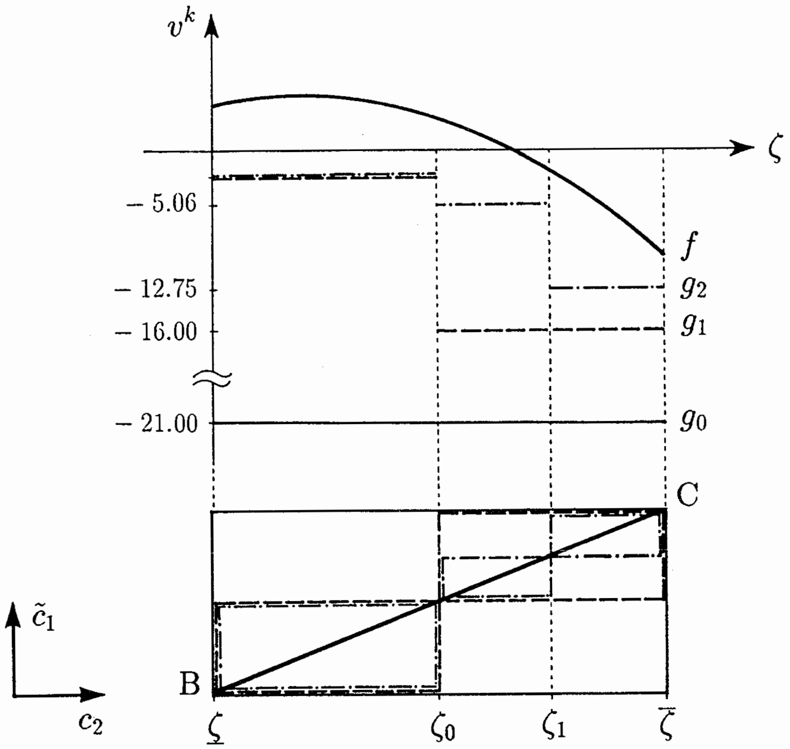

Figure 3.1. Illustration of BBP.

then terminate.

4° Let $\underline{Z}_j = [\underline{\zeta}_j, \zeta_j]$ and $\overline{Z}_j = [\zeta_j, \overline{\zeta}_j]$. Compute $v^k(\underline{Z}_j)$, $L_k(\underline{Z}_j)$, $v^k(\overline{Z}_j)$ and $L_k(\overline{Z}_j)$.

5° Let $\mathcal{Z} = \mathcal{Z} \cup \{\underline{Z}_j, \overline{Z}_j\}$. Let $j = j + 1$, and go to 2°. ◻

Theorem 3.3. *Procedure BBP terminates after finitely many iterations if $\epsilon > 0$. If $\epsilon = 0$ and BBP does not terminate, it generates an infinite sequence of points $x^k(\zeta_j)$'s, every accumulation point of which is a globally optimal solution of $(P_k(Z))$.*

Proof: Suppose the procedure does not terminate. Then an infinite sequence of intervals $Z_j = [\underline{\zeta}_j, \overline{\zeta}_j]$'s is generated in Z. We can take a subsequence Z_{j_ℓ}'s such that $(Z =)\ Z_{j_0} \supset Z_{j_1} \supset Z_{j_2} \supset \cdots$. Since Z_j is divided by the middle point $\zeta_j = (\underline{\zeta}_j + \overline{\zeta}_j)\,/\,2$, we can assume that $\overline{\zeta}_{j_\ell} - \underline{\zeta}_{j_\ell} = 2(\overline{\zeta}_{j_{\ell+1}} - \underline{\zeta}_{j_{\ell+1}})$ for every ℓ. Hence we have

$$\|x^k(\overline{\zeta}_{j_\ell}) - x^k(\underline{\zeta}_{j_\ell})\| = \|x^k(\underline{\zeta}) - x^k(\overline{\zeta})\| \,/\, 2^\ell \tag{3.11}$$

by linearity of $x^k(\zeta)$ over $D_k(Z)$.

Now we assume that there exists some positive constant σ such that

$$f(x^*) - v^k(Z_{j_\ell}) \geq \sigma, \ \ \forall \ell. \tag{3.12}$$

By continuity of f there is some positive value $\delta(\sigma)$ such that if

$$\|x' - x''\| < \delta(\sigma), \tag{3.13}$$

then $|f(x') - f(x'')| < \sigma$. It follows from (3.11) that (3.13) holds for any x', $x'' \in D_k(Z_{j_\ell})$ when ℓ is beyond a number:

$$\tilde{\ell}(\sigma) = \ln \|x^k(\overline{\zeta}) - x^k(\underline{\zeta})\| - \ln \delta(\sigma).$$

Moreover, we can see from (3.2) – (3.5) and (3.9) that if $x' \in L_k(\mathrm{Z}_{j_\ell})$, i.e., $f(x') = v^k(\mathrm{Z}_{j_\ell})$, then $\|x' - x\| < \delta(\sigma)$ for any $x \in D_k(\mathrm{Z}_{j_\ell})$. Therefore we have $f(x^k(\zeta_{j_\ell})) - v^k(\mathrm{Z}_{j_\ell}) < \sigma$ for $\ell > \tilde{\ell}(\sigma)$, which contradicts assumption (3.12). If $\epsilon > 0$, then (3.10) holds after finitely many iterations and BBP terminates.

Suppose $\epsilon = 0$. Then we have $\lim_{\ell \to \infty}(f(x^k(\zeta_{j_\ell})) - v^k(\mathrm{Z}_{j_\ell})) = 0$. Since we choose Z_{j_ℓ} with the least $v^k(\mathrm{Z}_{j_\ell})$ from $\mathcal{Z}$, we obtain

$$\lim_{\ell \to \infty} f(x^k(\zeta_{j_\ell})) = \lim_{\ell \to \infty} v^k(\mathrm{Z}_{j_\ell}) \leq f(x), \quad \forall x \in D_k(\mathrm{Z}). \qquad \square$$

To save the memory needed by BBP we can employ the depth first rule in choosing Z_j from $\mathcal{Z}$ instead of the best bound rule. Although the convergence is somewhat slower, this modification causes no trouble if $\epsilon > 0$. However, if $\epsilon = 0$, the sequence $x^k(\zeta_j)$'s might converge to some locally but not globally optimal solution of $(\mathrm{P}_k(\mathrm{Z}))$.

4. Computational Experiment

We will report the results of computational experiment on algorithm PRSM incorporating procedure BBP. We solved the following two subclasses of (P):

$$\left|\begin{array}{ll} \text{minimize} & (c_1{}^T x - c_{10})^2 - (c_1{}^T x - c_{10})(c_2{}^T x - c_{20}) \\ \text{subject to} & Ax \leq b, \ x \geq 0, \\ & c_k{}^T x \geq c_{k0}, \ k = 1, 2, \end{array}\right. \qquad (4.1)$$

$$\left|\begin{array}{ll} \text{minimize} & (c_1{}^T x - c_{10})^2 - (c_1{}^T x - c_{10}) \exp(c_{20} - c_2{}^T x) \\ \text{subject to} & Ax \leq b, \ x \geq 0, \\ & c_k{}^T x \geq c_{k0}, \ k = 1, 2, \end{array}\right. \qquad (4.2)$$

where $c_k \in \mathbb{R}^n$, $c_{k0} \in \mathbb{R}^1$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. All data of examples were randomly generated between -1.000 and 1.000. Problem (4.1) is a so-called linear multiplicative program, whose objective function can be expressed by the product of two affine functions, say $c_1{}^T x - c_{10}$ and $(c_1 - c_2)^T x - c_{10} + c_{20}$. If the product is quasiconcave on the feasible region, we can solve the problem efficiently by using the algorithms proposed in [9, 10, 13]. Unfortunately, the objective function of (4.1) is neither convex nor quasiconcave because $(c_1 - c_2)^T x - c_{10} + c_{20}$ can have both positive and negative values on the feasible region (see e.g. [9]). Hence the available algorithms do not work for (4.1).

In procedure BBP we employed the depth first rule in choosing Z_j from $\mathcal{Z}$. Also, among two subintervals $\underline{\mathrm{Z}}_j$ and $\overline{\mathrm{Z}}_j$ of Z_j we took out the one giving the less lower bound from $\mathcal{Z}$ before the other. The program was coded in C language and tested on a SUN SPARCstation ELC computer (20.5 mips).

Table 4.1 shows the computational results when the tolerance is fixed at $\epsilon = 10^{-5}$ and the size of problems ranges from $(m, n) = (200, 150)$ to $(350, 300)$. It contains the average number of pivot operations (including primal ones for the linear program solved in Step 1 of PRSM), branching operations and the average CPU time in seconds (and also their respective standard deviations in the brackets) needed for solving ten examples. Note that both problems (4.1) and (4.2) require the same number of pivot operations because their feasible regions are identical. Table 4.2 shows the results when (m, n) is fixed at $(200, 150)$ and ϵ ranges from 10^{-3} to 10^{-9}. The average number of branching operations and CPU time of ten examples are listed in it.

Table 4.1. Computational results when $\epsilon = 10^{-5}$.

m	200	200	250	250	300	300	350
n	150	200	200	250	250	300	300
Total number of pivots.							
	226.4 (30.124)	362.5 (96.225)	385.8 (98.238)	352.5 (90.155)	385.1 (120.45)	463.3 (160.515)	452.1 (209.555)
Total number of branchings.							
(4.1):	138.7 (78.411)	182.8 (100.152)	160.8 (83.244)	152.7 (98.105)	195.9 (82.440)	200.2 (122.173)	153.0 (102.326)
(4.2):	151.3 (82.002)	118.2 (94.448)	163.0 (94.043)	189.3 (140.644)	134.6 (94.291)	217.3 (106.39)	145.7 (121.598)
CPU time in seconds.							
(4.1):	46.040 (5.897)	83.130 (24.615)	124.515 (26.276)	117.942 (30.304)	174.678 (55.837)	233.958 (66.787)	279.792 (132.336)
(4.2):	46.305 (6.310)	82.972 (24.795)	124.525 (26.236)	118.463 (30.646)	173.562 (54.174)	234.020 (66.834)	279.613 (132.489)

We see from Tables 4.1 and 4.2 that algorithm PRSM can solve fairly large scale problems of both the classes (4.1) and (4.2) with enough accuracy when they are randomly generated. There is not much difference in the results between the two classes. It should be noted that the number of branching operations depends only upon the tolerance ϵ but not upon the size of (m, n). However, since the branching involves no hard operations such as a simplex pivot, it has a little influence on the computational time as shown in Table 4.2. The total computational time is consequently dominated by the number of iterations of the parametric simplex algorithm. Also its variance is reasonably small compared with the usual global optimization algorithms using cutting planes.

5. Average Performance of the Algorithm for Some Instances

As shown in Section 1, problem (P) involves numerous subclasses. Among them are the following two nonconvex quadratic programs:

Table 4.2. Computational results when $(m, n) = (200, 150)$.

ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}
Total number of branchings.				
(4.1):	36.4 (24.577)	138.7 (78.411)	255.9 (139.543)	373.7 (203.150)
(4.2):	34.4 (18.597)	151.3 (82.002)	286.7 (141.017)	423.7 (200.655)
CPU time in seconds.				
(4.1):	45.778 (6.140)	46.040 (5.897)	47.403 (6.298)	47.537 (6.739)
(4.2):	45.767 (6.212)	46.305 (6.310)	47.413 (6.092)	47.650 (6.587)

$$\text{(P1)} \quad \left| \begin{array}{ll} \text{minimize} & f_1(x) = (c_1{}^T x - c_{10})(c_2{}^T x - c_{20}) \\ \text{subject to} & x \in D, \end{array} \right. \tag{5.1}$$

$$\text{(P2)} \quad \left| \begin{array}{ll} \text{minimize} & f_2(x) = c_1{}^T x - (c_2{}^T x - c_{20})^2 \\ \text{subject to} & x \in D, \end{array} \right. \tag{5.2}$$

where $c_k \in \mathbb{R}^n$, $c_{k0} \in \mathbb{R}^1$ $(k = 1, 2)$ and $D \subset \mathbb{R}^n$ defined by (2.4). Linear multiplicative programs (P1) appear in many applications such as microeconomics [6], bond portfolio optimization [8], and computational geometry [11, 14] and so forth (see [10, 18]). If every feasible solution $x \in D$ satisfies $c_k{}^T x \geq c_{k0}$ for $k = 1, 2$, then (P1) is a quasiconcave minimization [9] and can be solved by the algorithms proposed in [9, 13] as well. Problem (P2) is a concave quadratic program, whose objective function f_2 has only one negative eigenvalue. In their recent article [19] Pardalos and Vavasis have proved the NP-hardness of (P2) by converting a clique problem on a graph to (5.2). Quadratic programs are known to be in NP [26], and hence (P2) is a NP-complete problem.

Here we will discuss the average performance of algorithm PRSM when we apply it to those nonconvex quadratic programs (P1) and (P2).

Recall that $(\mathrm{P}_k(\mathrm{Z}))$ solved by BBP is a minimization of f over the line segment $D_k(\mathrm{Z})$. If f is a quadratic function such as f_1 and f_2, we can calculate a rigorous solution of $(\mathrm{P}_k(\mathrm{Z}))$ analytically without calling procedure BBP. Hence the total number of arithmetic operations needed for solving (P1) and (P2) can be bounded only by that of dual pivot operations. Moreover, we can solve them even if the feasible region D is unbounded. In this case the parametric right-hand-side simplex algorithm would generate a basis $\underline{B}$ which is optimal to $(\mathrm{PL}_k(\zeta))$ $(k = 1, 2)$ for all $\zeta \in \mathrm{Z}' = [\underline{\zeta}, +\infty)$ for some $\underline{\zeta}$. At the same time it generates some direction vector $d \in \mathbb{R}^n$, and we have

$$D_k(\mathrm{Z}') = \{x \in \mathbb{R}^n \mid x = x^k(\underline{\zeta}) + \lambda d, \ \lambda \geq 0\}.$$

It is easy to check whether f_1 (f_2) is bounded from below on $D_k(\mathrm{Z}')$. If we find it unbounded, the original problem has no globally optimal solutions.

Let us again consider the linear programs $(\mathrm{PL}_k(\zeta))$, $k = 1, 2$. Denote by $g_k(\zeta)$ the objective function value of $(\mathrm{PL}_k(\zeta))$, i.e.,

$$g_1(\zeta) = \min\{c_2{}^T x \mid x \in D, \ c_1{}^T x = \zeta\}; \quad g_2(\zeta) = \max\{c_2{}^T x \mid x \in D, \ c_1{}^T x = \zeta\}.$$

Lemma 5.1. *Let $\zeta_{\inf} = \inf\{c_1{}^T x \mid x \in D\}$ and $\zeta_{\sup} = \sup\{c_1{}^T x \mid x \in D\}$. Then, (i) function g_1 is piecewise linear convex on the interval $(\zeta_{\inf}, \zeta_{\sup})$, (ii) function g_2 is piecewise linear concave on the interval $(\zeta_{\inf}, \zeta_{\sup})$.*

Proof: Follows from a well-know result on linear programming (see e.g. [2]). □

We can regard PRSM as a method which generates the analytic form of g_k and compute a global minimum of f over the line segment corresponding to each linear piece of g_k. Under Assumption 2.1 the number of linear pieces of g_k's coincides with that of dual pivot operations of PRSM.

If we take the partial dual with respect to the constraint $c_1{}^T x = \zeta$ of $(\mathrm{PL}_1(\zeta))$, then

$$\begin{aligned} g_1(\zeta) &= \min_x \{c_2{}^T x \mid x \in D, \ c_1^T x = \zeta\} \\ &= \min_x \sup_{\eta \in \mathbb{R}^1} \{c_2{}^T x + \eta(c_1{}^T x - \zeta) \mid x \in D\} \\ &= \sup_{\eta \in \mathbb{R}^1} \{-\eta\zeta + \min\{\eta c_1{}^T x + c_2{}^T x \mid x \in D\}\}. \end{aligned}$$

Letting

$$h_1(\eta) = \min\{\eta {c_1}^T x + {c_2}^T x \mid x \in D\},$$

we have

$$g_1(\zeta) = \sup_{\eta \in \mathbb{R}^1} \{-\eta\zeta + h_1(\eta)\}. \tag{5.3}$$

Similarly, g_2 can be reduce to

$$g_2(\zeta) = \inf_{\eta \in \mathbb{R}^1} \{-\eta\zeta + h_2(\eta)\}, \tag{5.4}$$

where

$$h_2(\eta) = \max\{\eta {c_1}^T x + {c_2}^T x \mid x \in D\}.$$

The following lemma is analogous to Lemma 5.1:

Lemma 5.2. *(i) Function h_1 is piecewise linear concave on $\mathbb{R}^1$, (ii) function h_2 is piecewise linear convex on $\mathbb{R}^1$.* □

Thus we can see from (5.3) and (5.4) that if the analytic form of h_k is given, we can obtain that of g_k in $O(I_k)$ time, where I_k represents the number of linear pieces of h_k. The number of linear pieces of g_k is obviously $O(I_k)$.

Adler and Haimovich have proved in [1, 5] that the average number of linear pieces I_k is bounded by $O(\min\{m, n\})$ under sign-invariant probabilistic assumptions imposed on the data (A, b, c_1, c_2). (Readers are referred to an excellent survey article by Shamir [21] or a book by Schrijver [20] for the results proved in the unpublished manuscripts [1, 5].) In their probabilistic model, Assumption 2.1 is fulfilled with probability one. This implies that the average number of dual pivot operations required by PRSM is also bounded by $O(\min\{m, n\})$. On the other hand, the linear program to be solved in Step 1 of PRSM is a standard linear program, which is well known to be solved in polynomial time. Consequently, the average number of arithmetic operations needed for solving (P1) and (P2) is lower-order polynomial relative to the size of A. A similar result for a certain class of bilinear programs has been proved in [12].

The key of the above discussion is the polynomial solvability of (P_k(Z)). If f is quasi-concave on D, either of the extreme points $x^k(\underline{\zeta})$ and $x^k(\overline{\zeta})$ of $D_k(\mathrm{Z})$ is optimal to (P_k(Z)). Hence we can also solve such instances of (P) in polynomial time on the average.

Acknowledgment

The author is grateful to anonymous referees for their helpful comments and suggestions.

References

[1] Adler, I., "The expected number of pivots needed to solve parametric linear programs and the efficiency of the self-dual simplex method," Draft, Department of IEOR, University of California, Berkeley (CA, 1983).

[2] Chvătal, V., *Linear Programming*, Freeman and Company (NY, 1983).

[3] Gal, T., "Linear parametric programming – a brief survey," *Mathematical Programming Study* **21** (1984), 43 – 68.

[4] Geoffrion, M., "Solving bicriterion mathematical programs," *Operations Research* **15** (1967), 39 – 54.

[5] Haimovich, M., "The simplex method is very good! – On the expected number of pivot steps and related properties of random linear programs," Draft, Uris Hall, Columbia University (NY, 1983).

[6] Henderson, J. M. and R. E. Quandt, *Microeconomic Theory*, Mcgraw-Hill (NY, 1971).

[7] Horst, R. and H. Tuy, *Global Optimization: Deterministic Approaches*, Springer-Verlag (Berlin, 1990).

[8] Konno, H. and M. Inori, "Bond portfolio optimization by bilinear fractional programming," *Journal of the Operations Research Society of Japan* **32** (1988), 143 – 158.

[9] Konno, H. and T. Kuno, "Linear multiplicative programming," *Mathematical Programming* **56** (1992), 51 – 64.

[10] Konno, H. and T. Kuno, "Multiplicative programming problems," in R. Horst and P.M. Pardalos (eds.), *Handbook of Global Optimization*, Kluwer Academic Publishers (Dortrecht, 1995).

[11] Konno, H., T. Kuno, S. Suzuki, P. T. Thach and Y. Yajima, "Global optimization techniques for a problem in the plane," Report IHSS91-36, Institute of Human and Social Sciences, Tokyo Institute of Technology (Tokyo, 1991).

[12] Konno, H., T. Kuno and Y. Yajima, "Parametric simplex algorithms for a class of NP complete problems whose average number of steps are polynomial," *Computational Optimization and Applications* **1** (1992), 227 – 239.

[13] Konno, H., Y. Yajima and T. Matsui, "Parametric simplex algorithms for solving a special class of nonconvex minimization problems," *Journal of Global Optimization* **1** (1991), 65 – 82.

[14] Kuno, T., "Globally determining a minimum-area rectangle enclosing the projection of a higher-dimensional set," *Operations Research Letters* **13** (1993), 295 – 303.

[15] Kuno, T., and Y. Yamamoto, "A parametric simplex algorithm for solving a class of rank-two reverse convex programs," Report ISE-TR-93-103, Institute of Information Sciences and Electronics, University of Tsukuba (Tsukuba, 1993).

[16] Muu, L. D., "An algorithm for solving convex programs with an additional convex-concave constraint," *Mathematical Programming* **61** (1993), 75 – 87.

[17] Muu, L. D. and W. Oettli, "Method for minimizing a convex-concave function over a convex set," *Journal of Optimization Theory and Application* **70** (1991), 377 – 384.

[18] Pardalos, P. M., "Polynomial time algorithms for some classes of constrained nonconvex quadratic problems," *Optimization* **21** (1990), 843 – 853.

[19] Pardalos, P. M. and S. A. Vavasis, "Quadratic programming with one negative eigenvalue is NP-hard," *Journal of Global Optimization* **1** (1991), 15 – 22.

[20] Schrijver, A., *Theory of Linear and Integer Programming*, John-Wiley & Sons (NY, 1986).

[21] Shamir, R., "The efficiency of the simplex method: a survey," *Management Science* **33** (1987), 301 – 334.

[22] Sniedovich, M., Macalalag, E. and S. Findlay, "The simplex method as a global optimizer: a c-programming perspective," *Journal of Global Optimization* **4** (1994), 89 – 109.

[23] Swarup, K., "Indefinite quadratic programming," *Cahiers du Centre d'Études de Recherche Opérationnelle* **8** (1966), 217 – 222.

[24] Tuy, H., "The complementary convex structure in global optimization," *Journal of Global Optimization* **2** (1992), 21 – 40.

[25] Tuy, H., S. Ghannadan, A. Migdalas and P. Värbrand, "Strongly polynomial algorithm for a production-transportation problem with concave production costs," *Optimization*

27 (1993), 205 – 227.
[26] Vavasis, S. A., "Quadratic programming is in NP," *Information Processing Letters* **36** (1990), 73 – 77.

Takahito KUNO:
Institute of Information Sciences and Electronics
University of Tsukuba
Tsukuba, Ibaraki 305, Japan
(E-mail: takahito@is.tsukuba.ac.jp)